

RESEARCH

Open Access



# Genome-wide identification and expression analysis of the *SPL* transcription factor family and its response to abiotic stress in *Pisum sativum* L

Long Li<sup>1,3</sup>, Jian bo Xu<sup>2</sup>, Zhi wen Zhu<sup>2</sup>, Rui Ma<sup>4</sup>, Xiao zong Wu<sup>2,5\*</sup> and Yu ke Geng<sup>1\*</sup>

## Abstract

Squamous promoter binding protein-like (*SPL*) genes encode plant-specific transcription factors (TFs) that play essential roles in modulating plant growth, development, and stress response. Pea (*Pisum sativum* L.) is a coarse grain crop of great importance in food production, biodiversity conservation and molecular genetic research, providing genetic information and nutritional resources for improving agricultural production and promoting human health. However, only limited researches on the structure and functions of *SPL* genes exist in pea (*PsSPLs*). In this study, we identified 22 *PsSPLs* and conducted a genome-wide analysis of their physical characteristics, chromosome distribution, gene structure, phylogenetic evolution and gene expression patterns. As a result, the *PsSPLs* were unevenly distributed on the seven chromosomes of pea and harbored the SBP domain, which is composed of approximately 76 amino acid residues. The phylogenetic analysis revealed that the *PsSPLs* clustered into eight subfamilies and showed high homology with *SPL* genes in soybean. Further analysis showed the presence of segmental duplications in the *PsSPLs*. The expression patterns of 22 *PsSPLs* at different tissues, developmental stages and under various stimulus conditions were evaluated by qRT-PCR method. It was found that the expression patterns of *PsSPLs* from the same subfamily were similar in different tissues, the transcripts of most *PsSPLs* reached the maximum peak value at 14 days after anthesis in the pod. Abiotic stresses can cause significantly up-regulated *PsSPL19* expression with spatiotemporal specificity, in addition, four plant hormones can cause the up-regulated expression of most *PsSPLs* including *PsSPL19* in a time-dependent manner. Therefore, *PsSPL19* could be a key candidate gene for signal transduction during pea growth and development, pod formation, abiotic stress and plant hormone response. Our findings should provide insights for the elucidating of development regulation mechanism and breeding for resistance to abiotic stress pea.

**Keywords** *Pisum sativum* L, *SPL* genes, Genome-wide analysis, Growth and development, Abiotic stress, Plant hormone response, *PsSPL19*

\*Correspondence:

Xiao zong Wu  
wuxzong@zzuli.edu.cn

Yu ke Geng  
gengyuke0502@163.com

<sup>1</sup>Minzu University of China, 100010 Beijing, P.R. China

<sup>2</sup>School of Food and Biological engineering, Zhengzhou University of Light Industry, 450002 Zhengzhou, P.R. China

<sup>3</sup>College of Agronomy, Hebei Agricultural University, 071001 Baoding, P.R. China

<sup>4</sup>State Key Laboratory of Crop Stress Biology for Arid Areas, College of Plant Protection, Northwest A&F University, 712100 Yangling, Shaanxi, P.R. China

<sup>5</sup>Zhengzhou University of Light Industry, 450002 Zhengzhou, P.R. China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Pea (*Pisum sativum*L.) is a legume plant and a member of the genus *Pisum*. It is an annual climbing plant and an important source of food [1, 2]. *Pisum sativum* is a source of high-quality proteins and other nutrients such as dietary fiber, vitamin C, vitamin K, and minerals [3, 4]. Additionally, pea is used to produce green manure, which improves soil quality by fixing nitrogen and providing nutrients to crops [5, 6]. Meanwhile, the strong drought - and cold - resistant properties made pea adaptable to a wide range of climates [7].

Transcription factors (TFs) are a class of proteins that regulate gene transcription. They regulate gene expression by binding to specific regions in gene promoters, activating or inhibiting gene transcription, and thus regulate a variety of biological processes, such as plant growth, development, differentiation and response to environmental stimuli [8, 9]. Squamosa promoter binding protein-like (SPL) TFs are a ubiquitous family of TFs in various plants and contain a conserved SBP domain, two Cys2/His2 type zinc fingers, and two types of zinc fingers which is a structurally stable zinc finger motif formed by two cysteines (Cys) and two histidines (His) via ring and helical connections, and this domain binds to specific regions of DNA [10–12]. SPL is usually involved in plant growth, flower organ development and fruit development [10, 13]. In order to explore the biological functions and regulatory mechanisms of *SPLs* in plant growth, development and stress response, various plant species have been studied and elaborated by scholars, such as *Arabidopsis thaliana* [14, 15], *Morus alba* [16], *Nicotiana tabacum* [17], *Zea mays* [18], *Setaria italica* [19], *Vaccinium corymbosum* [20], *Vitis vinifera* [21], *Fagopyrum tataricum* [22], *Solanum lycopersicum* [23], *Codonopsis pilosula* [23], *Triticum aestivum* [24], and *Jatropha curcas* [25].

SPL TFs were first identified in Snapdragon (*Antirrhinum majus* L.). AmSPL regulates the formation and development of flower organs via binding to promoters of genes related to flower development [26]. AtSPL9 and AtSPL15 in *Arabidopsis thaliana* are highly expressed in stem apex meristem and participate in leaf differentiation to regulate plant leaf morphology and structure [27, 28]. Cui et al. confirmed that AtSPL9 is negatively regulated by miR156, which promotes early flowering of *Arabidopsis thaliana* under salt and drought stress. In addition, AtSPL9 can compete with TT8 (Transparent testa8) combined with PRODUCTION OF ANTHOCYANIN PIGMENT 1 (PAP1) [29], such competition interferes with the stability of MYB-bHLH-WD40 (MBW) transcription-activating complex, directly prevents the expression of flavonoid synthesis gene dihydroflavonol 4-reductase (DFR), and thus negatively regulates anthocyanin accumulation [15]. AtSPL3 and AtSPL4 are significantly differentially expressed in the apical meristem of

*Arabidopsis thaliana*, but both are involved in the regulation of flowering [15, 30]. Loss-of-function mutants in *OsSPL14*, *OsSPL16* and *OsSPL18* lead to flower dysplasia, including increased inflorescence branching, atypical flower morphology and anther development defects [31]. This implies that these SPL TFs play essential regulatory roles in the formation and development of rice flower organs.

In this study, we identified *PsSPLs* from pea genome sequence obtained from publicly available databases using bioinformatics analysis. A genome-level comprehensive analysis was performed to explore their structural characteristics, motif composition, chromosome localization and evolutionary relationships. Then, we used real-time quantitative fluorescent PCR (qRT-PCR) to analyze the expression patterns of *PsSPLs* of different tissues, expression changes during pod formation stage, induced expression level under abiotic stress and hormone response, separately. The results provide an experimental basis for further revealing the mechanism of pea *SPL* gene family in regulating the growth and development, yield and environmental response, which possesses important application value for molecular breeding of pea with high yield and quality.

## Methods

### Gene identification

The whole-genome sequence of *Pisum sativum* was retrieved from the Ensembl Genomes website (<http://ensemblgenomes.org/>), whereas the whole-genome sequence of *Arabidopsis thaliana* was downloaded from The *Arabidopsis* Information Resource (TAIR) website (<https://www.arabidopsis.org/>). BLASTp [32] (with a score cutoff of  $\geq 100$  and an e-value cutoff of  $\leq 1e-10$ ) was used to align *AtSPL* with the complete genome sequence of closely related species rye to identify candidate *SPL* genes. Subsequently, the hidden Markov model (HMM) of the SBP conserved domain was retrieved from the Pfam database (<http://pfam.sanger.ac.uk/>), and HMMER 3 (<http://plants.ensembl>) was used to identify non-redundant *SPL* genes [33]. The selected *SPL* genes were confirmed using the SMART tool (<http://smart.embl-heidelberg.de/>) [34, 35]. A total of 22 *SPL* genes were obtained for the subsequent analyses. ExPasy tool was used to determine the physical characteristics of the genes, including sequence length, isoelectric point (pI), and molecular weight (MW). The WoLFPSORT tool (<https://wolffpsort.hgc.jp/>) was used to predict the subcellular localization of the selected *SPL* genes.

### Gene structure, chromosomal distribution and gene duplication analysis and classification of the *PsSPLs*

The MEME website (<http://meme.nbcr.net/meme/intro.html>) was used to identify the motifs in *PsSPLs*, with a

maximum of 10 motifs set for the search. The GSDS (<http://gsds.cbi.pku.edu.cn>) tool was used to analyze the distribution of introns and exons in the sequences of *PsSPLs*. The Circos website was used to identify the location of *PsSPLs* in chromosomes. MCScanX toolkit was used to evaluate the collinear duplication events on the chromosomes [36]. Mega7.0 software was used to perform multiple sequence comparison of *SPL* genes from pea and *Arabidopsis thaliana*, and the phylogenetic analysis of different *SPL* genes was performed by neighborhood linkage (NJ) method and phylogenetic tree construction [37]. At the same time, the NJ method was also used for phylogenetic analysis of *SPL* genes from Pea, *Oryza sativa*, *Zea mays*, *Hordeum vulgare*, *Arabidopsis thaliana*, *Triticum aestivum* (hexaploid) and *Aegilops tauschii* Coss. Moreover, the evolutionary relationships among different species and subfamilies are further discussed.

#### Evaluation of varying growth conditions and the effect of different hormone treatments on *Pisum sativum*

In this study, pea No. 6 cultivar provided by Zhengzhou University of Light Industry was taken as the research material. Fully mature pea seeds were selected and evenly sown in a greenhouse under the following growth conditions: 16 h light at 28 °C / 8 h darkness at 20 °C, and 75% constant relative humidity. The Tissues (roots, stems, leaves, anthers and styles, seeds) of pea plants with the same growth state at 21-day growth are collected and stored in an ultra-low temperature refrigerator of -80 °C. Parallely, Pea plants were treated with four abiotic stresses, including salt stress (5% sodium chloride), low temperature stress (4 °C), drought stress (30% PEG 6000) and high temperature stress (40 °C) for 0, 1, 4 and 12 h, respectively, and the untreated plants obtained at the same time point were used as controls, the whole plants were collected and stored in an ultra-low temperature. After that, the pea plants at the filling stage were treated with ABA, IAA, JA, SA and GA3 at concentrations of 100 μM for 0, 1, 4 and 12 h, respectively and the whole plants were sampled and preserved at ultra-low temperature. In the above experiments, three plants were selected for one type of treatment, and three independent biological replicates were set up for each.

#### Total RNA extraction, cDNA synthesis, and qRT-PCR analysis

RNA extraction kit (Tiangen Biotech, Beijing, China) was used to extract RNA from different tissue samples collected. HiScript Q RT Super Mix for qPCR (Vazyme #R122) was used for reverse transcription to obtain cDNA, and AceQ qPCR SYBR Green Master Mix was used for qRT-PCR analysis of *SPL* gene. Primers for *SPL* gene qRT-PCR experiments were designed using

Primer5.0 software (Additional file 5: Table S5). In this study, the glyceraldehyde-3-phosphate dehydrogenase (*GADPH*) gene was used as an internal reference for gene expression quantification. The experiment was repeated three times and the relative gene expression was calculated by  $2^{-\Delta\Delta Ct}$  method. Compared with the control group, we defined a significant difference in expression levels as  $\geq 2$ -fold or  $\leq 0.5$ -fold [38].

#### Statistical analysis

In this study, all data were subjected to analysis of variance (ANOVA) using the JMP 6.0 software (SAS Institute). The least significant difference (LSD) test was performed to determine the significance of differences between different treatment groups [38, 39]. A significance level (e.g., 0.05 or 0.01) is typically set in the LSD test to determine if the differences are statistically significant. Origin version 8.0 (OriginLab, Northampton, MA, USA) was used to generate the histograms for gene expression levels.

## Results

### Identification of *SPL* genes in Pea

A comprehensive and systematic analysis of the *PsSPLs* was conducted, leading to identification and characterization of 22 *PsSPLs*. The length of proteins encoded by these genes ranged from 139 (*PsSPL3*) to 1025 (*PsSPL9*) amino acids. The maximum molecular weight of the proteins encoded by *PsSPL* was 113.918 kDa (*PsSPL18*), whereas the minimum molecular weight was 15.9 kDa (*PsSPL4*). The pI values ranged from 6 to 9.28 (*PsSPL7* and *PsSPL12*), with an average value of 7.622. All the *PsSPLs* harbored the conserved SBP domain, and four of them contain ANK (ankyrin) domain (Additional file 1: Table S1). The ANK domain is a conserved domain involved in protein-protein interaction and it plays an important role in various biological processes, including cell signaling, protein interactions and assembly of protein complexes. Subcellular localization prediction revealed that all the *PsSPLs* were localized in the nucleus, however, some were observed in different cellular tissues with 9 *SPLs* localized in the endoplasmic reticulum, 6 *SPLs* in the chloroplasts and seven in the cytoplasm (Additional file 1: Table S1).

### *PsSPLs* exhibit high evolutionary conservation

Phylogenetic analysis of 22 *PsSPLs* from pea and 16 *AtSPLs* from *Arabidopsis thaliana* was conducted in this study. The *PsSPLs* formed three branches (Group 1–8) based on the classification method proposed by Cenci and Rouard consistent with the classification of *SPL* proteins in *Arabidopsis thaliana* [40]. This finding indicated conservation of *SPLs* across the evolutionary process without any loss events. In addition, we observed

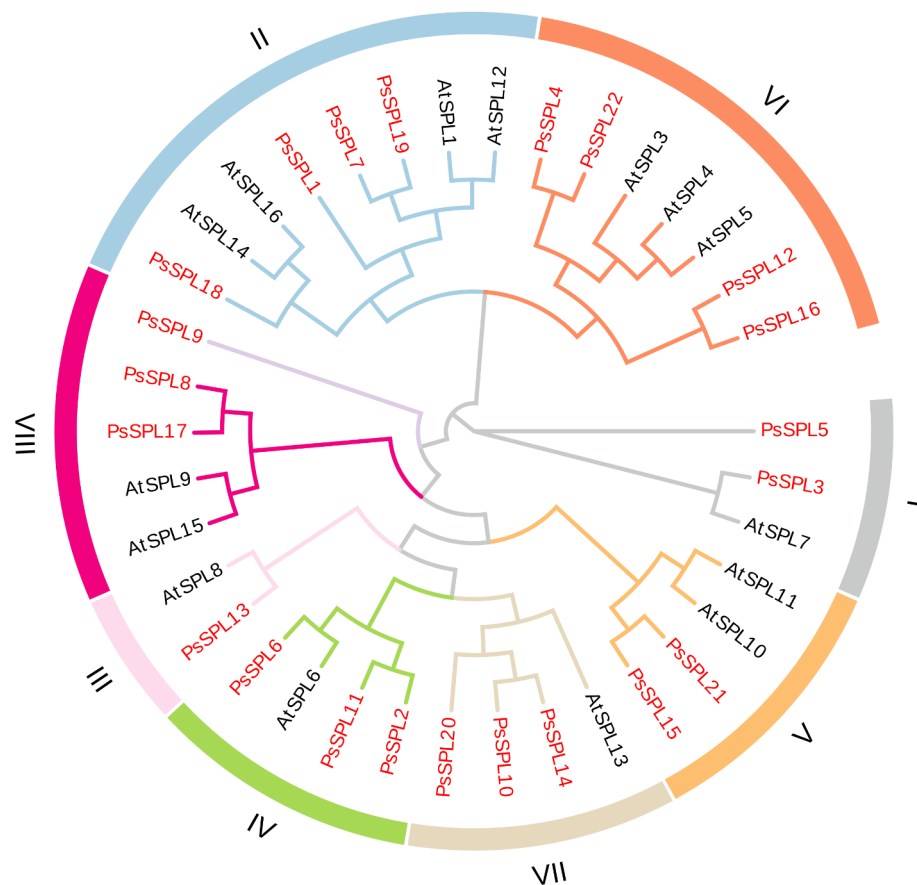
an uneven distribution of *SPLs* exist among the eight subfamilies. Subfamilies II and VI contained the largest number of members (4 *PsSPLs*), whereas subfamily III possessed the lowest number of genes (only 1 *PsSPL*) (Fig. 1). Subfamilies VII, VIII, and IV owned the same number of *SPLs* (3 *PsSPLs*). Subfamilies I and V each contain 2 *PsSPLs* (Fig. 1; Additional file 1: Table S1). Comparative analysis of the phylogenetic tree revealed that *PsSPL* clustered with *AtSPL* (bootstrap support  $\geq 70$ ), implying that these proteins may exert similar biological functions. These findings would provide a basis for further exploring the evolution and function of the *SPL* gene family in plants.

### Conserved motifs and structure analysis of the *PsSPLs*

Comparing the exons and introns of genes provided a comprehensive understanding of gene structure, function, and regulation in an organism [41, 42]. The results showed that different genes exhibited varying numbers of exons and introns, ranging from 0 to 14. *PsSPL19* exhibited the highest number of introns (14), whereas *PsSPL5* lacked intronic structures (Fig. 2). *PsSPL1*, *PsSPL7*, *PsSPL18*, and *PsSPL19*, members of subfamily

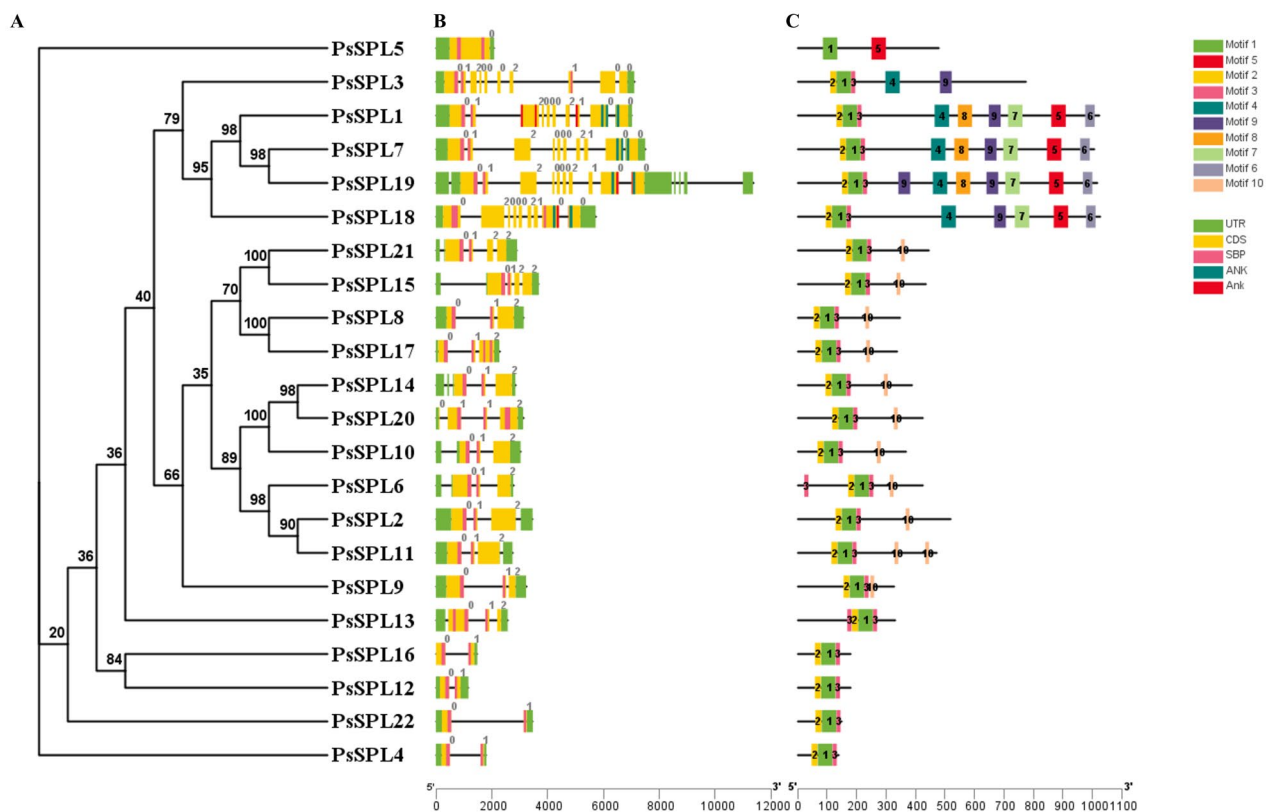
II, exhibited the most complex gene structures, indicating that subfamily II may have different functions. Typically, members of the same gene family exhibit conserved intron and exon structures. This conservation can be attributed to the conserved structure of introns and exons caused by gene replication and recombination during the evolution of gene family members.

We identified 10 conserved motifs conserved across the genomic sequences of different genes using the MEME tool [43, 44], indicating the homology and conservation of regions in these genes. We named these motifs as Motif 1–10. The different subfamilies exhibited some differences in motifs, but all genes had Motif 1, indicating that Motif 1 is a conserved domain in *SPL* genes (Fig. 2). Subfamily II exhibited 9 of the 10 motifs identified in this study (lacked Motif 10). All genes exhibited Motif 1, Motif 2, and Motif 3. Subfamilies IV, V, VII, and VIII had Motif 10, which was absent in the other subfamilies. Motif 2 was located at the beginning of all patterns, whereas Motif 1 was positioned between Motif 2 and Motif 3 (Fig. 2; Additional file 2: Table S2). The conservation of motifs implied that the subfamily potentially emerged through evolutionary processes such as gene



**Fig. 1** Phylogenetic analysis of *SPL* proteins in pea and *Arabidopsis thaliana*. 22 *PsSPLs* were divided into eight clades (I–VIII) and indicated with different colors. The red pentacle represents Pea, the black circle represents *Arabidopsis thaliana*





**Fig. 2** Phylogenetic relationships, gene structure, and motif distributions of *PsSPL* genes. Phylogenetic tree was constructed for each node with 1000 replicates using the NJ method (A). Exons and introns are indicated by yellow rectangles and grey lines, respectively (B). These numbers indicate the annotation file of the different phases of gene CDS, which are defined as "0", "1", and "2". C Amino acid motifs in the SPL proteins (1–10) were represented by colored boxes, the black lines indicate relative protein lengths

duplication and recombination. These findings indicate that the genes may form clusters in the genome with similar structural and functional features.

#### Chromosomal distribution and gene duplication of *PsSPLs*

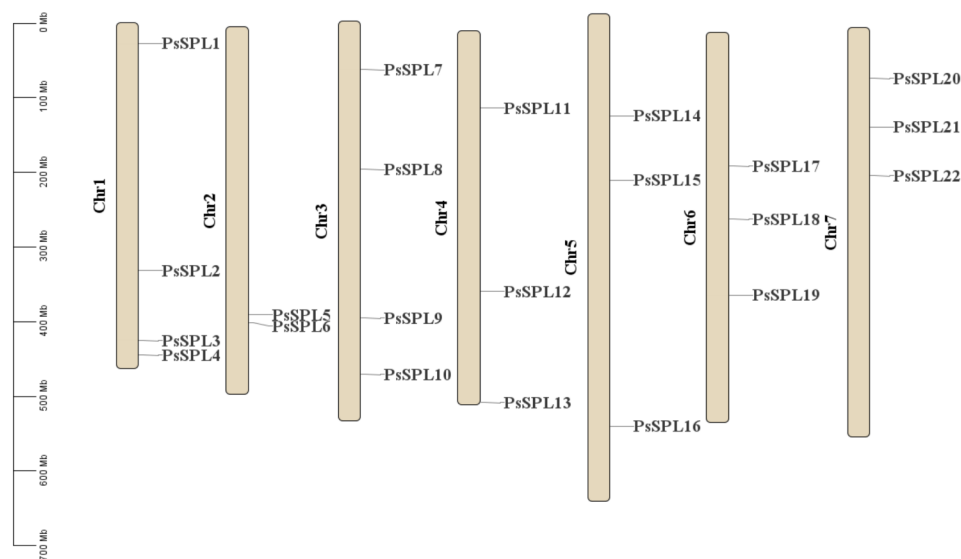
The *SPL* genes were mapped to different chromosomes using the latest version of the Pea genome database [45]. The 22 *PsSPLs* were unevenly distributed across seven chromosomes (Fig. 3). Chr1 and Chr3 harbored the highest number of *PsSPLs* (4 genes, approximately 18.18%), followed by Chr4, Chr5, Chr6, and Chr7, which each contained three *SPLs* (approximately 13.63% each). Chr2 harbored the least number of *PsSPLs* (2 genes, approximately 9.09%). The *PsSPLs* are randomly distributed across different chromosomes with no significant correlation between the distribution of *SPLs* on chromosomes to their function and structure.

Tandem duplication events and segmental duplication events are key forms of gene duplication processes, which contribute to genomic diversity and play important roles in shaping the structure and function of the genome [46, 47]. In this study, no any tandem duplication event was detected on the chromosomes (Fig. 4; Additional

file 3: Table S3). However, we found eight homologous sequences produced by four pairs of fragments distributed on five chromosomes, this indicated the existence of evolutionary relationships among *PsSPLs*, Chr5 contains the largest number of *PsSPLs* ( $n=3$ ), followed by Chr 7 ( $n=2$ ), while Chr 3, Chr 4, and Chr 6 each contains only one *PsSPL* (Fig. 4).

#### *PsSPLs* are evolutionarily related to *SPLs* in other species

We compared the *PsSPL* gene with *SPL* genes from two dicot plants (*Arabidopsis thaliana* and *Glycine max*) and four monocot plants (*Setaria italica*, *Triticum aestivum*, *Oryza sativa* and *Chenopodium quinoa*) to evaluate the genomic diversity among different species and explore the common ancestors and evolutionary relationships. We constructed a phylogenetic tree using the NJ method based on 10 conserved motifs identified in the genes, obvious distinctions in motifs among the different species were observed, but the same subfamily exhibited motif conservation (Fig. 5). Most genes harbored Motif 1, Motif 2, and Motif 4. Subfamily II exhibited all the 10 motifs, implying that this subfamily may have a more complex and diverse gene regulatory network compared



**Fig. 3** Schematic representation of the chromosomal distribution of *PsSPLs*. Vertical bars represent the chromosomes of *Pea*. The chromosome number is indicated to the left of each chromosome. The scale on the left represents chromosome length

with the other subfamilies. All members of the subfamily except for subfamily VI started with Motif7. Most subfamilies exhibited high homology with the *GmSPL* gene family of soybean, indicating that pea and soybean may share a common ancestor or had high degree of structural and sequence conservation during evolution.

We conducted a collinearity analysis based on the genes of *Pisum sativum* and six representative species to elucidate the evolutionary relationships and functional conservation across different species. The results revealed collinearity between 22 *PsSPL* genes and genes from *Arabidopsis thaliana* (15), *Setaria italica* (18), *Chenopodium quinoa* (23), *Glycine max* (29), *Oryza sativa* (29), and *Triticum aestivum* (56). Monocot plants and pea exhibited a closer evolutionary relationship than the relationship between dicot and monocot plants. We identified 52 gene pairs shared between *Glycine max* and *Pisum sativum*, 15 gene pairs shared with *Arabidopsis thaliana*, 7 gene pairs shared with *Oryza sativa*, 7 gene pairs shared with *Setaria italica*, 6 gene pairs shared with *Triticum aestivum*, and 6 gene pairs shared with *Chenopodium quinoa*.

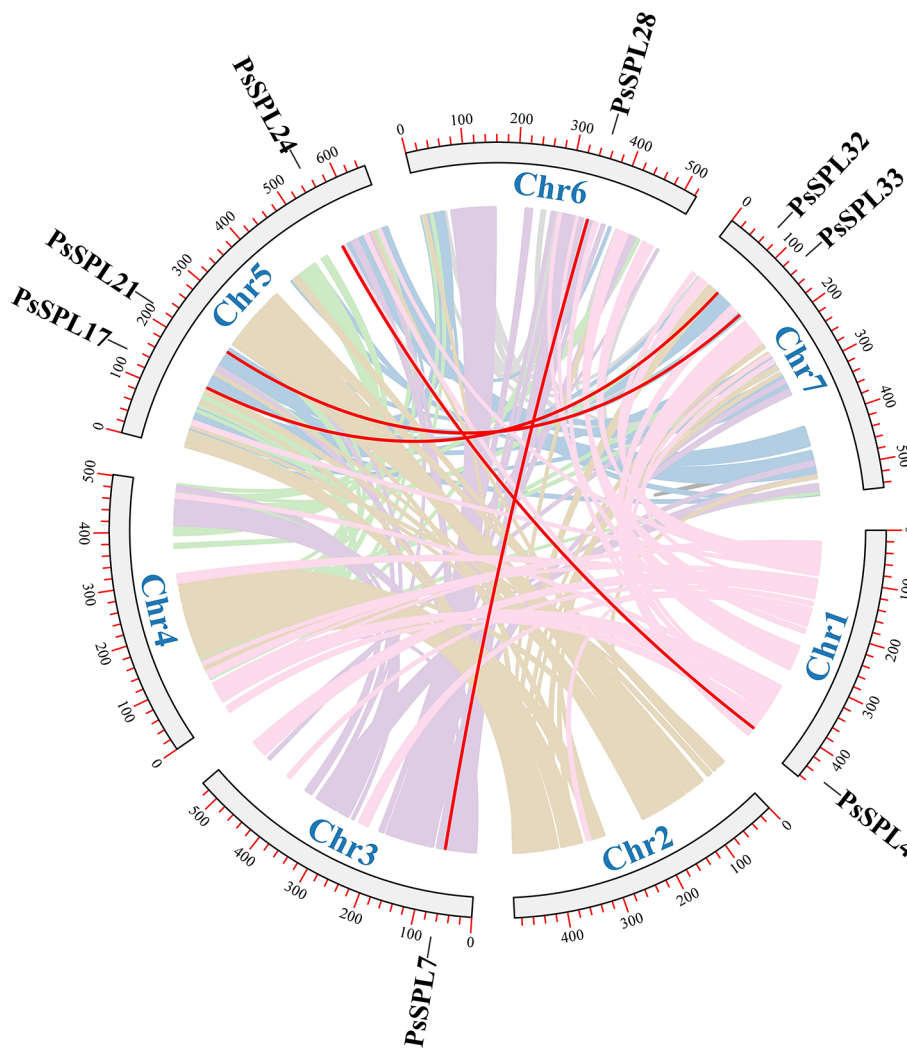
Comparative analysis of collinearity among the six plants revealed that the number of collinear *SPL* genes between *Pisum sativum* and *Triticum aestivum* and *Chenopodium quinoa* was the lowest, whereas the number of collinear *SPL* genes between *Pisum sativum* and *Glycine max* was the highest. Interestingly, these species shared some common collinear genes. For example, *PsSPL3* exhibited collinearity with AT5G18830.3 / Os05t0408200-01 / TraesCS1A02G255300.1 / KQL15519. This finding indicates that these collinear genes are highly conserved and possibly existed before

the divergence of these species (Fig. 6; Additional file 4: Table S4). Collinear genes are used to explore genomic similarities and differences, evolutionary relationships, and functional conservation among different species. The findings on collinear genes provide information on genetic variations and evolutionary relationships among different species.

#### Expression patterns of the *PsSPLs* in different plant organs

In order to evaluate the potential function of *PsSPLs*, qRT-PCR was used to analyze the expression of 22 *PsSPLs* in four organs: root, stem, leaf, flower and fruit. The results show that *PsSPLs* were expressed differently in these four organs, which reflects the specific biological functions and adaptation strategies of plants at different growth and development stages and under different environmental conditions. Three genes (*PsSPL3*, *PsSPL7*, and *PsSPL19*) were most expressed in stems, while eight genes (*PsSPL2*, *PsSPL5*, *PsSPL6*, *PsSPL9*, *PsSPL11*, *PsSPL14*, *PsSPL15*, and *PsSPL20*) were most expressed in leaves. *PsSPL1*, *PsSPL12*, *PsSPL18* and *PsSPL20* were highly expressed in flowers (Fig. 7). Genes from the same subfamily may have maintained similar expression patterns over the course of evolution, reflecting that they may share a common ancestor and have maintained similar regulatory patterns over the course of evolution. It is obvious that the expression levels of all *PsSPLs* in roots are lower than those in stems, leaves and flowers, so we speculate that *SPLs* may be closely related to the development of stems, leaves and flowers of plants.

Previous studies have shown that the expression level and pattern of *SPL* during fruit development may be related to biological processes such as fruit growth,



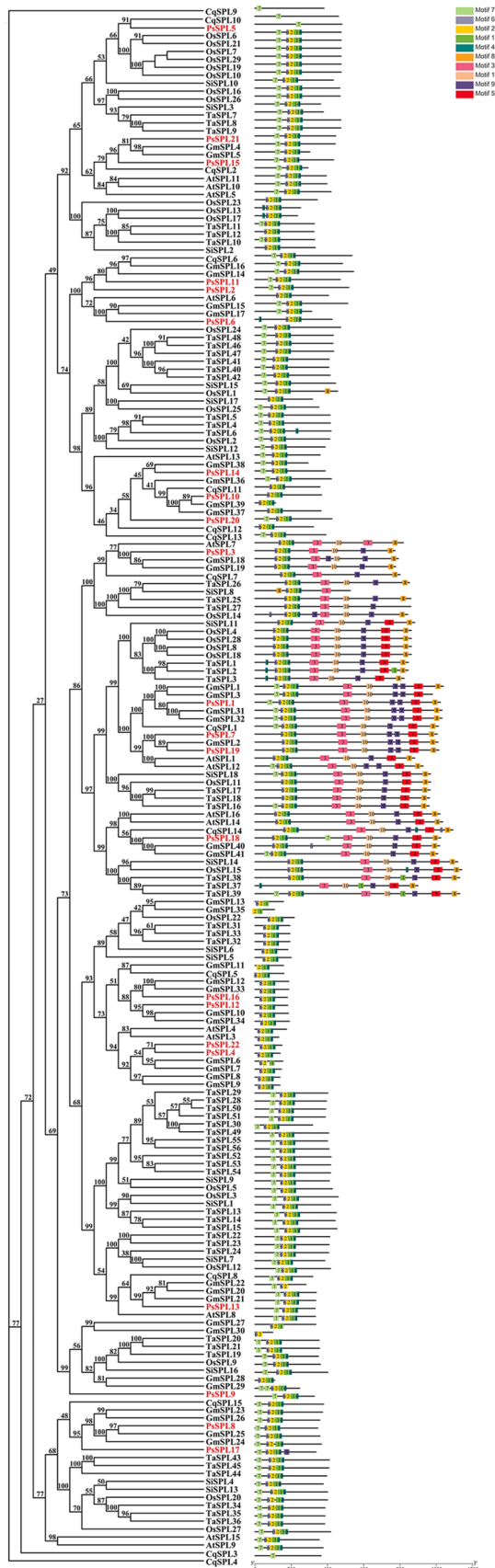
**Fig. 4** Schematic representation of the chromosomal distribution and interchromosomal relationships of *PsSPLs*. Colored lines denote all synteny blocks in the Pea genome, and the red lines denote duplicated *SPL* pairs. The chromosome number is denoted at the bottom of each chromosome

ripening and storage. Therefore, we selected 15 representative genes for qRT-PCR verification at five post-anthesis stages (7DPA, 14DPA, 21DPA, 28DPA and 35DPA). As shown in Fig. 7, almost all genes showed differential expression patterns at different time periods. In pea fruits, we found that *PsSPL4* had the most significant differential expression at 14DPA stage, while *PsSPL3* showed a downward trend with the increase of time. Interestingly, *PsSPL9* remained stable over time, suggesting that it may not be associated with fruit development. In addition, we found that most of the significantly differentially expressed genes (DEGs) were concentrated at 14DPA stage, such as *PsSPL2*, *PsSPL4*, *PsSPL6*, *PsSPL8*, *PsSPL11*, *PsSPL12*, *PsSPL13*, *PsSPL14*, *PsSPL15*, *PsSPL17*, *PsSPL2*, *PsSPL4*, *PsSPL6*, *PsSPL8*, *PsSPL19* and *PsSPL22*. These results indicated that the expression of most *PsSPLs* were mainly concentrated in the early stage of pea fruit. The expression level of *PsSPL1* reached its maximum value at

35DPA stage, which may be related to the later stage of fruit development. In addition, we found that the expression of *PsSPL7* was the highest in the pod, and most of the genes showed significant differential expression. The expression of *PsSPL5* and *PsSPL16* was not significantly different at different times, indicating that it was not related to pod development of pea. The heat map showed that there was a correlation between different *PsSPLs*, and most *PsSPLs* were positively correlated. However, some *PsSPLs* were negatively correlated, such as *PsSPL6* and *PsSPL21/PsSPL1*, and *PsSPL1* and *PsSPL9* ( $P < 0.05$ ).

#### Expression patterns of *PsSPLs* under various abiotic stress conditions

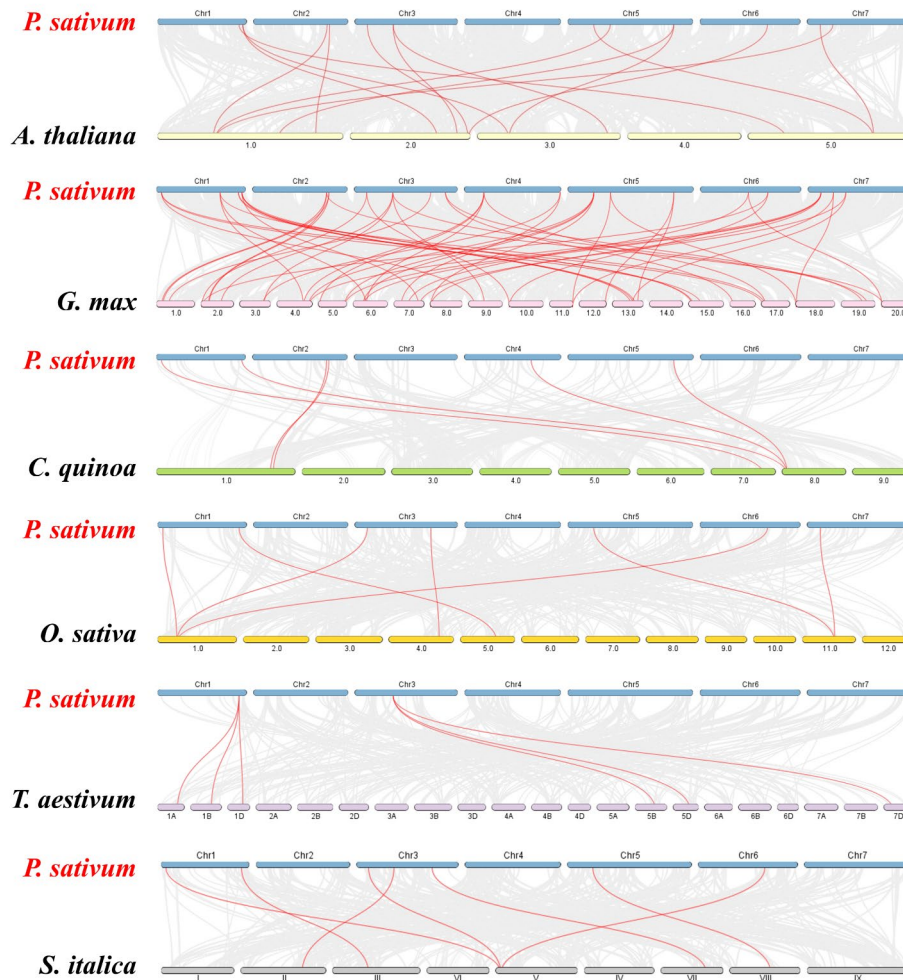
Prediction of gene function by exposing plants to various stress treatments can offer valuable insights into elucidating the biological functions of genes under adverse conditions, providing information on plant adaptation



**Fig. 5** Phylogenetic relationships and motif compositions of the PsSPLs with six different plant species (*Arabidopsis thaliana*, *Solanum lycopersicum*, *Vitis vinifera*, *Sorghum bicolor* Moench, *Oryza sativa* and *Zea mays*). Outer panel: an unrooted phylogenetic tree constructed using Geneious R11 with the neighbor-joining method. Inner panel: distribution of conserved motifs in SPL proteins. The differently colored boxes represent different motifs and their positions in each SPL protein sequence. The sequence information for each motif is provided in Additional file 2 (Table S2)

mechanisms and regulatory networks in response to stress [48]. We used four stress treatments to explore the effects of various stress conditions on the expression patterns of *PsSPL* genes in different tissues. The results showed that various stress conditions induced differential expression of *PsSPL* genes in different tissues and at different time points. *PsSPL* genes are implicated in stress signaling pathways, so their expression levels are upregulated with increase in treatment time, indicating their important regulatory role in alleviating stress in plants. The expression of most *SPL* genes was upregulated in stems under cold stress conditions. *PsSPL1*, *PsSPL3*, *PsSPL7*, *PsSPL10*, *PsSPL18*, and *PsSPL19* genes exhibited significant differential expression in roots, leaves, and stems, with upregulation at the early stages of stress treatment and then downregulation at the late stages. *PsSPL19* had the highest expression level under PEG and salt stress condition, and the expression of the other genes was also upregulated under these two stress treatments. Conversely, the expression of *PsSPL5*, *PsSPL15*, *PsSPL16*, *PsSPL17*, and *PsSPL20* were downregulated under PEG and salt stress conditions. Tissue-specific responses were mainly observed in stems under cold and heat stress treatments. *PsSPL4*, *PsSPL10*, *PsSPL13*, *PsSPL14* and *PsSPL18* genes displayed similar expression patterns, with increases expression levels in different tissues observed over treatment time. On the contrary, the expression of *PsSPL5*, *PsSPL16* and *PsSPL20* was downregulated over time. The genes showed different expression patterns under different treatments, and the significant downregulation observed at the start of stress induction may be attributed to the rapid regulation and adaptation process of genes to the treatments. *PsSPL19* expression was upregulated under the four stress treatments in different tissues, indicating its potentially important role in stress responses. We performed correlation analysis and generated a heatmap to show the correlation in expression levels among *PsSPL* members under stress. The results showed that the expression levels of *PsSPL2*, *PsSPL3*, *PsSPL6*, *PsSPL9*, *PsSPL11*, *PsSPL12*, *PsSPL14*, *PsSPL15*, *PsSPL18*, and *PsSPL19* genes were significantly positively correlated, whereas the expression of *PsSPL1* was significantly positively correlated with *PsSPL5* expression (Fig. 8). The expression of some *PsSPL* genes was significantly negatively correlated with the expression of *PsSPL5* and *PsSPL20*.





**Fig. 6** Synteny analyses of the *SPL* genes between *Pisum sativum* and six representative plant species (*Arabidopsis thaliana*, *Glycine max*, *Chenopodium quinoa*, *Oryza sativa*, *Triticum aestivum* and *Setaria italica*). Gray lines on the background indicate the collinear blocks in *Pisum sativum* and other plant genomes; red lines highlight the syntenic of *SPL* gene pairs in *Pisum sativum*

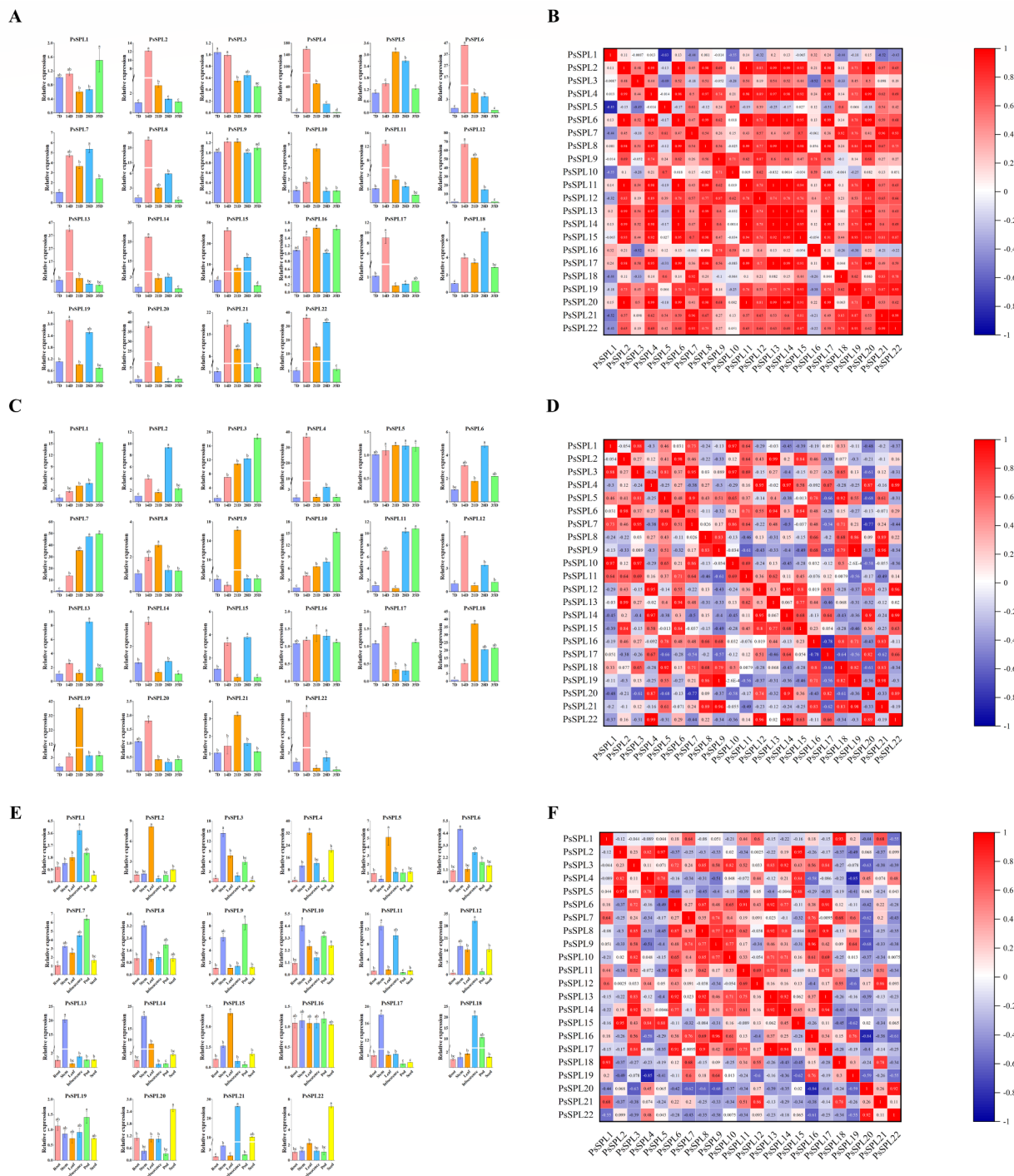
Previous studies demonstrated that some *SPL* genes are involved in hormone regulation [49–51]. Therefore, we treated pea plants with five hormones (ABA, JA, SA, GA, and IAA) and evaluated the changes in the transcription levels *SPL* genes. The expression of most genes was upregulated under ABA, JA, SA, and IAA treatments, whereas most genes were downregulated under GA treatment (Fig. 9). *PsSPL1* showed the highest expression level under ABA and GA induction, whereas *PsSPL13* exhibited the highest expression level under SA and JA induction. *PsSPL19* had the highest expression level under IAA induction. Interestingly, these five hormones exhibited peak expression levels at different times. For example, most *PsSPLs* showed the highest expression levels at 12 h under ABA induction, 1 h under SA induction, and 4 h under JA induction. This finding implies that different hormones regulate the growth and development of

pea plants by modulating the expression levels of *PsSPLs* at different times. We observed upregulation in the expression of *PsSPL5* under SA treatment, whereas the expression of this gene was downregulated under other hormone treatments.

## Discussion

Pea is a vegetable crop and an important legume crop [1], which is rich in proteins, vitamins and minerals and is widely used in food processing and household consumption [2]. In agriculture production, peas are commonly used as green manure or feed plants to improve soil quality and can be rotated with other crops to increase agricultural yields [52]. Additionally, the pea genome is relatively small and the genetic background is relatively simple, so it is easy to study and manipulate [53, 54]. Therefore, peas are widely used as a model plant for plant



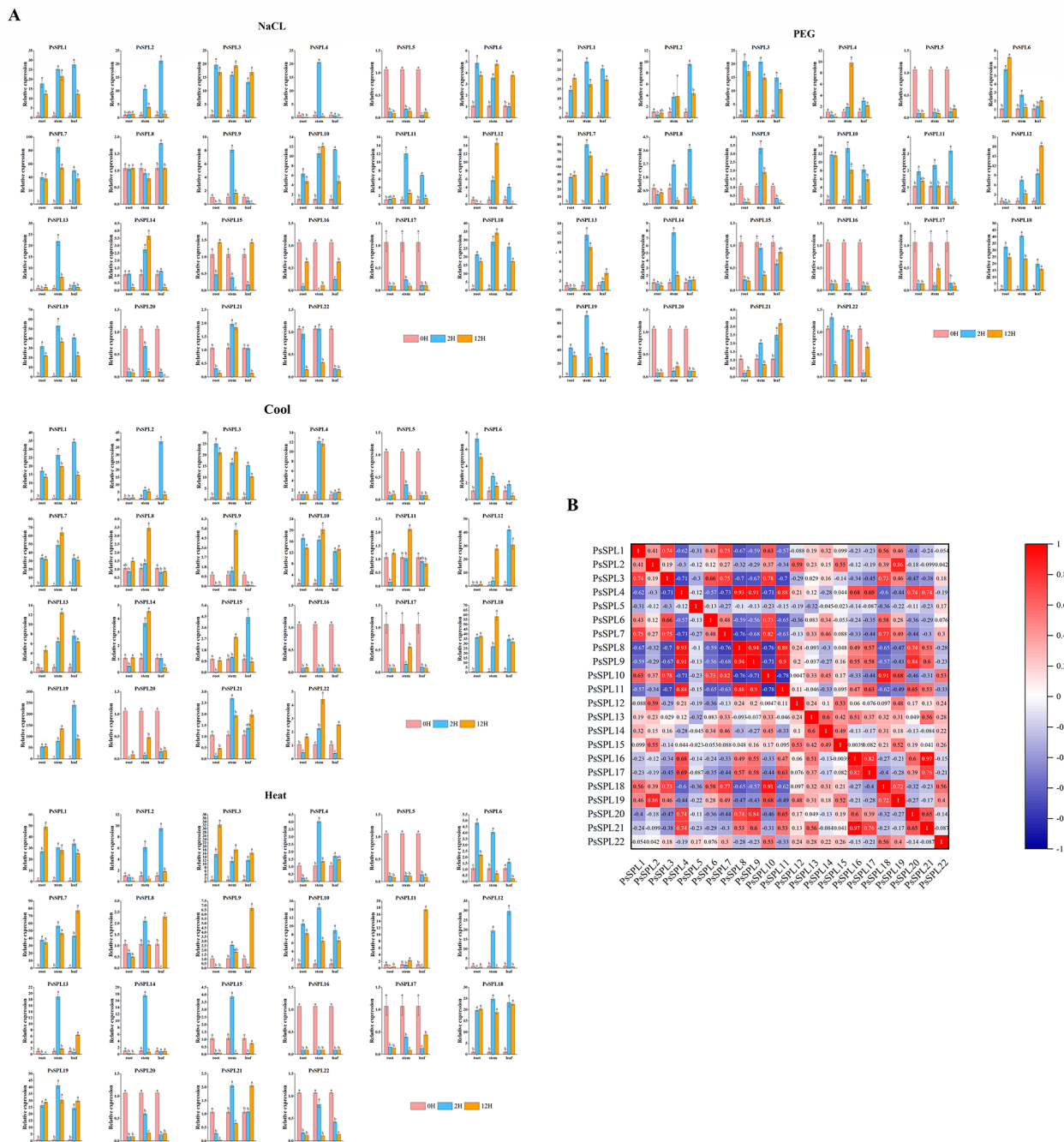


**Fig. 7** Tissue-specific gene expression of 22 *PsSPLs* and gene expression during fruit development. Expression patterns of 22 *PsSPLs* and in fruit of different stage and flower, leaf, root, stem were detected by qPCR. Error bars are obtained from three measurements. Lowercase letters indicate significant differences between treatments ( $\alpha=0.05$ , LSD) (A, C, E). Coexpression analysis of 22 *PsSPLs* (B, D, F)

genetics, biology and agricultural research. The *SPL* gene family plays an important role in fruit development [55–57]. Investigating the function and regulation mechanism of *SPL* gene family members during pea fruit development and abiotic stresses can make better understand the

molecular mechanisms underlying pea pod formation. Furthermore, it provides a crucial theoretical basis for improving crop varieties and fruit yield and quality.

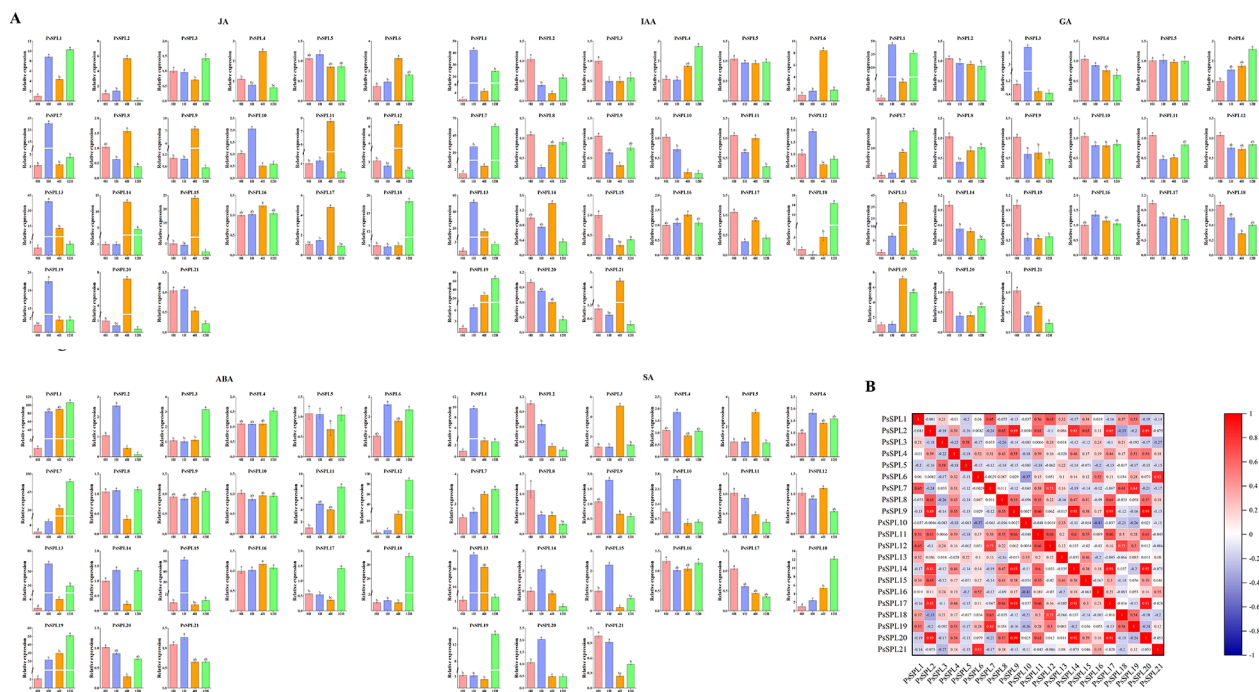
We identified 22 *SPLs* in the pea genome and explored the sequence and structural characteristics of these



**Fig. 8** Expression of *PsSPLs* in plants subjected to abiotic stresses (PEG, NaCl, heat, and cool treatments) at the seedling stage in three organs (root, stem, and leaf). **(A)** Changes in expression of representative genes analyzed by qRT-PCR. Error bars were obtained from three measurements. The lowercase letter above the bar indicates a significant difference ( $\alpha=0.05$ , LSD) among the treatments. **(B)** Coexpression analysis of 22 *PsSPLs* in several plant organs. Positive numbers: positive correlations; negative numbers: negative correlations. Red numbers indicate a significant correlation at the 0.05 level

genes. The SPL Protein sizes range from 139 (*PsSPL3*) to 1025 (*PsSPL9*) amino acids. The maximum molecular weight of these proteins was 113.918 kDa (*PsSPL18*) and the minimum molecular weight was 15.9 kDa (*PsSPL4*). The number of introns in all *PsSPLs* ranges from 0 to 14. *PsSPL19* has a maximum of 14 introns, while *PsSPL5* has no intron structure. Introns play a key role in genome

evolution by regulating gene diversity, gene expression regulation, and influencing protein diversity. The intron structure of subfamily II genes is the most complex, suggesting that subfamily II genes may have more functions than other genes. Intron-free genes usually exhibit higher transcription and translation efficiency than intron-containing genes. It is speculated that under



**Fig. 9** Expression analysis of 22 *PsSPLs* in fruits in response to different hormones (JA, ABA, IAA, SA and GA). **(A)** qRT-PCR was utilized to detect the expression patterns of 22 *PsSPLs*. Error bars ( $n=3$ ) represent the standard error. Lowercase letters above the bars indicate significant differences ( $\alpha=0.05$ , LSD) among treatments. **(B)** Coexpression analysis of 22 *PsSPLs*. Positive numbers = positive correlation; negative numbers = negative correlation. Red numbers indicate a significant correlation at the 0.05 level

stress conditions, the transcriptional regulation process of intron-free genes may be simpler and faster after responding to stimuli [58–60].

In this study, 22 *PsSPLs* and their homologous in *Arabidopsis thaliana* were clustered. The *SPLs* of both were found to be divided into eight subfamilies, each containing at least one *AtSPL* gene (Fig. 1). Obvious structural and functional differences may be structural and functional differences exist among different *SPL* subfamilies. There are four fragment repetition events in *PsSPL*, which may be generated during the evolution of *PsSPL* family [46, 47]. The diversity of sequence length within the *PsSPL* family may confer adaptability to different environmental and ecological conditions on peas, thus affecting the survival of pea species [61]. The *SPLs* clustering results of pea and other six plants showed that *PsSPLs* were highly homologous to soybean *SPLs* (Fig. 5), although there were obvious differences in morphology and growth habits between the two crops. Phylogenetic analysis shows that subfamily II members are very complex and contain all identified conserved motifs. As a binding protein to gene cis-elements, *SPL* plays a role as a key regulatory element [62]. The diversity of motifs may cause *PsSPL* to bind promoter cis-elements of multiple genes to regulate downstream gene expression in a complex way [63, 64], thus exerting diverse functions.

Gene expression analysis is often used for functional prediction to elucidate regulatory mechanisms and biological processes [65]. qRT-PCR analysis showed that most representative *PsSPLs* were highly expressed in stems, leaves and flowers, and were involved in the regulation of pod development in pea (Fig. 7). Rice *OsSPL4*, *OsSPL13* and *OsSPL16* promote the increase of rice grain width by regulating the level of cytokinin synthesis, and ultimately affect the grain morphological size [66–68], which seems to support the involvement of *SPLs* in seed or fruit formation. Under the four non-stress conditions, most of the genes of pea were induced to up-regulate, and the expressions of *PsSPL1*, *PsSPL5*, *PsSPL7* and *PsSPL18* of subfamily II were significantly up-regulated in stems under cold and heat stress (Fig. 8). It has been reported that *AtSPL1* and *AtSPL12*, members of II subfamily enhanced the heat resistance of inflorescence by regulating ABA signaling pathway, thereby reducing the sensitivity of flower organs to high temperature stress [69], suggesting that some *SPLs* may have similar regulatory mechanisms in the response of pea and *Arabidopsis thaliana* to heat stress. Moreover, miR-156*SPL* can target *AtSPL9*, regulate the binding activity of *AtSPL9* with C-REPEAT binding factor 2(CBF2) and induce the expression of *AtSPL9* under cold stress, and enhance freezing resistance [70]. In apples, the miR156/*MdSPL13* module regulates the salt tolerance of apples

by targeting the promoter of *MdWRKY100*, and overexpression of *MdWRKY100* will enhance the salt tolerance of apples [71], indicating that miR156 usually mediates the response of *SPL* to salt stress and drought stress. In addition, OsSPL10 imparts drought resistance to rice by regulating the expression of *OsNAC2* [72], while *TaSPL6* plays a negative regulatory role in plant drought stress response by reducing the expression of some genes involved in stress response, and overexpression leads to increased sensitivity of wheat to drought stress [73]. In addition, we predict the cis-acting elements of the *PsSPLs* promoters in order to find transcription factors that can bind to them, which may be closely related to abiotic stress response (Additional file 6: Table S6). All the above studies provide research ideas for revealing the mechanism of *PsSPL* regulating abiotic stress response. Under four different plant hormone treatments, the expression of most *SPLs* was increased to varying degrees, mainly at 1 h, 4 h and 12 h after SA, JA and ABA hormone treatments, and the expression of *PsSPL1* was increased under ABA and GA induction (Fig. 9). Previous studies have shown that in rice, OsSPL12 can directly interact with nine proteins related to GA signal and participate in the regulation of GA synthesis in rice grains [49], providing clues and references for the mechanism by which *PsSPL* family participates in GA signal transduction pathway to regulate pod development, promote seed dormancy and inhibit ear germination, etc. *PsSPL13* is highly expressed by SA and JA. *PsSPL19* may be associated with disease resistance, while IAA induced high expression of *PSSPL19* which may promote pod cell elongation or fruit enlargement.

## Conclusion

In this study, we identified 22 *PsSPLs* in the pea genome and conducted a comprehensive analysis of the structure and potential functions of these genes. The 22 *PsSPLs* were unevenly distributed across seven chromosomes and were classified into eight subfamilies based on homology with *AtSPLs*. All *PsSPLs* harbored the SBP domain. Subfamily II exhibited the most complex intron-exon structure and had the highest number of motifs, indicating the functional diversity of genes in this group. The *PsSPLs* family in peas lacked tandem duplications, but segmental duplications were observed, implying that segmental duplication was involved in the evolution of the gene family. Phylogenetic analysis showed that *PsSPLs* were highly homologous to soybean *SPLs*. The expression profile analysis of *PsSPLs* indicated that these genes may play an important role in the pea growth and development, pod maturation and environmental response. In particular, *PsSPL19* is considered as a potential candidate gene for exploring pea breeding.

## Abbreviations

SPL	SQUAMOSIA promoter-binding protein-like
<i>PsSPL</i>	<i>Pisum sativum</i> L. SPL
qRT-PCR	quantitative real-time polymerase chain reaction
AtSPL	<i>Arabidopsis thaliana</i> SPL
HMM	Hidden Markov Model
pI	isoelectric point
LG	linkage group
DPA	days post anthesis

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-024-10262-w>.

Supplementary Material 1  
 Supplementary Material 2  
 Supplementary Material 3  
 Supplementary Material 4  
 Supplementary Material 5  
 Supplementary Material 6

## Acknowledgements

Thanks to all our colleagues for providing useful discussions and technical assistance. Great thanks to the editor and reviewers for critically evaluating the manuscript and providing constructive comments for its improvement.

## Author contributions

LL: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing-original draft. JBX: Methodology, Software, Investigation, Writing-review & editing. ZWZ: Methodology, Validation. RM: Data curation, Formal analysis, Investigation. XZW: Conceptualization, Supervision. YKG: Funding acquisition, Conceptualization, Supervision. All authors reviewed the manuscript.

## Funding

This research was supported by the Fundamental Research Funds for the Central Universities (2022QNPY54).

## Data availability

Whole genome sequence information for *Pea* was obtained from the Ensembl genome website (<http://ensemblgenomes.org>). The seed used in this experiment was Zhongwan 6. The datasets supporting the conclusions of this study are included in the article and in additional files.

## Declarations

### Ethics approval and consent to participate

This article does not contain any studies involving human participants or animals performed by the authors. These methods were carried out in accordance with relevant guidelines and regulations. All experimental protocols were approved by Zhengzhou University of Light Industry.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that there are no competing interests.

Received: 15 January 2024 / Accepted: 27 March 2024

Published online: 31 May 2024

## References

- Pandey AK, Rubiales D, Wang Y, Fang P, Sun T, Liu N, Xu P. Omics resources and omics-enabled approaches for achieving high productivity and improved quality in pea (*Pisum sativum* L.). TAG theoretical and applied genetics Theoretische Und Angewandte Genetik 2021, 134(3):755–76.
- Teshome A, Bryngelsson T, Dagne K, Geleta M. Assessment of genetic diversity in Ethiopian field pea (*Pisum sativum* L.) accessions with newly developed EST-SSR markers. BMC Genet. 2015;16:102.
- Jarambasa T, Regon P, Jyoti SY, Gupta D, Panda SK, Tanti B. Genome-wide identification and expression analysis of the *Pisum sativum* (L.) APETALA2/ethylene-responsive factor (AP2/ERF) gene family reveals functions in drought and cold stresses. Genetica. 2023;151(3):225–39.
- Bo PT, Bai Y, Dong Y, Shi H, Soe Htet MN, Samoon HA, Zhang R, Tanveer SK, Hai J. Influence of Different Harvesting Stages and Cereals-Legume Mixture on Forage Biomass Yield, Nutritional Compositions, and Quality under Loess Plateau Region. Plants (Basel, Switzerland) 2022, 11(20).
- Singh T, Raturi H, Uniyal SP. Effect of Biofertilizer and Mulch on Growth, Yield, Quality and Economics of pea (*Pisum sativum* L.). Indian J Agricultural Res 2021.
- Cárcamo HA, Herle CE, Lupwayi NZ. *Sitona lineatus* (Coleoptera: Curculionidae) larval feeding on *Pisum sativum* L. affects Soil and Plant Nitrogen. J Insect Sci (Online) 2015, 15(1).
- Lahuta LB, Szablińska-Piernik J, Horbowicz M. Changes in metabolic profiles of pea (*Pisum sativum* L.) as a result of repeated short-term Soil Drought and subsequent re-watering. Int J Mol Sci 2022, 23(3).
- Feng K, Hou XL, Xing GM, Liu JX, Duan AQ, Xu ZS, Li MY, Zhuang J, Xiong AS. Advances in AP2/ERF super-family transcription factors in plant. Crit Rev Biotechnol. 2020;40(6):750–76.
- Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, et al. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Volume 290. New York, NY: Science; 2000. pp. 2105–10. 5499.
- Yamasaki K, Kigawa T, Inoue M, Tateno M, Yamasaki T, Yabuki T, Aoki M, Seki E, Matsuda T, Nunokawa E, et al. A novel zinc-binding Motif revealed by solution structures of DNA-binding domains of Arabidopsis SBP-family transcription factors. J Mol Biol. 2004;337(1):49–63.
- Klein J, Saedler H, Huijser P. A new family of DNA binding proteins includes putative transcriptional regulators of the Antirrhinum majus floral meristem identity gene SQUAMOSA. Mol Gen Genetics: MGG. 1996;250(1):7–16.
- Shao Y, Zhou HZ, Wu Y, Zhang H, Lin J, Jiang X, He Q, Zhu J, Li Y, Yu H, et al. OsSPL3, an SBP-Domain protein, regulates Crown Root Development in Rice. Plant Cell. 2019;31(6):1257–75.
- He B, Gao S, Lu H, Yan J, Li C, Ma M, Wang X, Chen X, Zhan Y, Zeng F. Genome-wide analysis and molecular dissection of the SPL gene family in *Fraxinus mandshurica*. BMC Plant Biol. 2022;22(1):451.
- Chen X, Zhang Z, Liu D, Zhang K, Li A, Mao L. SQUAMOSA promoter-binding protein-like transcription factors: star players for plant growth and development. J Integr Plant Biol. 2010;52(11):946–51.
- Usami T, Horiguchi G, Yano S, Tsukaya H. The more and smaller cells mutants of Arabidopsis thaliana identify novel roles for SQUAMOSA PROMOTER BINDING PROTEIN-LIKE genes in the control of heteroblasty. Development. 2009;136(6):955–64.
- Li H, Ma B, Luo Y, Wei W, Yuan J, Zhai C, He N. The Mulberry SPL Gene Family and the response of MnSPL7 to Silkworm Herbivory through activating the transcription of MnTT2L2 in the Catechin Biosynthesis Pathway. Int J Mol Sci 2022, 23(3).
- He L, Peng X, Cao H, Yang K, Xiang L, Li R, Zhang F, Liu W. The NTSPL Gene Family in *Nicotiana tabacum*: genome-wide investigation and expression analysis in response to cadmium stress. Genes 2023, 14(1).
- Mao H-D, Yu L-J, Li Z-J, Yan Y, Han R, Liu H, Ma M. Genome-wide analysis of the SPL family transcription factors and their responses to abiotic stresses in maize. Plant Gene. 2016;6:1–12.
- Lai D, Fan Y, Xue G, He A, Yang H, He C, Li Y, Ruan J, Yan J, Cheng J. Genome-wide identification and characterization of the SPL gene family and its expression in the various developmental stages and stress conditions in foxtail millet (*Setaria italica*). BMC Genomics. 2022;23(1):389.
- Feng X, Zhou B, Wu X, Wu H, Zhang S, Jiang Y, Wang Y, Zhang Y, Cao M, Guo B, et al. Molecular characterization of SPL gene family during flower morphogenesis and regulation in blueberry. BMC Plant Biol. 2023;23(1):40.
- Hou H, Li J, Gao M, Singer SD, Wang H, Mao L, Fei Z, Wang X. Genomic organization, phylogenetic comparison and differential expression of the SBP-box family genes in grape. PLoS ONE. 2013;8(3):e59358.
- Liu M, Sun W, Ma Z, Huang L, Wu Q, Tang Z, Bu T, Li C, Chen H. Genome-wide identification of the SPL gene family in Tartary Buckwheat (*Fagopyrum tataricum*) and expression analysis during fruit development stages. BMC Plant Biol. 2019;19(1):299.
- Yang J, Guo Z, Wang W, Cao X, Yang X. Genome-wide characterization of SPL Gene Family in *Codonopsis pilosula* reveals the functions of CpSPL2 and CpSPL10 in promoting the Accumulation of secondary metabolites and growth of *C. pilosula* Hairy Root. Genes 2021, 12(10).
- Zhu T, Liu Y, Ma L, Wang X, Zhang D, Han Y, Ding Q, Ma L. Genome-wide identification, phylogeny and expression analysis of the SPL gene family in wheat. BMC Plant Biol. 2020;20(1):420.
- Yu N, Yang JC, Yin GT, Li RS, Zou WT. Genome-wide characterization of the SPL gene family involved in the age development of *Jatropha curcas*. BMC Genomics. 2020;21(1):368.
- Ma J, Nie G, Yang Z, Ma S, Fan J, Hu R, Wu F, Zhang X. Genome-wide identification, characterization, and expression profiling analysis of SPL Gene Family during the Inflorescence Development in *Trifolium repens*. Genes 2022, 13(5).
- Wang LY, Yu P, Lyu JY, Hu YF, Han C, Bai MY, Fan M. BZR1 physically interacts with SPL9 to regulate the vegetative phase change and cell elongation in Arabidopsis. Int J Mol Sci. 2021;22(19):10415.
- Hu TQ, Manuela D, Xu ML. SQUAMOSA PROMOTER BINDING. PROTEIN-LIKE 9 and 13 repress BLADE-ON-PETIOLE 1 and 2 directly to promote adult leaf morphology in Arabidopsis. J Exp Bot. 2023;74(6):1926–39.
- Cui LG, Shan JX, Shi M, Gao JP, Lin HX. The miR156-SPL9-DFR pathway coordinates the relationship between development and abiotic stress tolerance in plants. Plant Journal: Cell Mol Biology. 2014;80(6):1108–17.
- Schwarz S, Grande AV, Bujdosó N, Saedler H, Huijser P. The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in Arabidopsis. Plant Mol Biol. 2008;67(1–2):183–95.
- Yao S, Yang Z, Yang R, Huang Y, Guo G, Kong X, Lan Y, Zhou T, Wang H, Wang W, et al. Transcriptional regulation of miR528 by OsSPL9 orchestrates antiviral response in Rice. Mol Plant. 2019;12(8):1114–22.
- Jacob A, Lancaster J, Buhler J, Harris B, Chamberlain RD. Mercury BLASTP: accelerating protein sequence alignment. ACM Trans Reconfigurable Technol Syst. 2008;1(2):9.
- Marchin M, Kelly PT, Fang J. Tracker: continuous HMMER and BLAST searching. Bioinf (Oxford England). 2005;21(3):388–9.
- Punta M, Coghill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, et al. The pfam protein families database. Nucleic Acids Res. 2012;40(Database issue):D290–301.
- Letunic I, Bork P. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 2018;46(D1):D493–6.
- Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49.
- Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. Mol Biol Evol. 2016;33(7):1870–4.
- Livak KJ. TD Schmittgen 2001 Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods (San Diego Calif) 25 4 402–8.
- Alaïda MF, Aldhebiani AY. Comparative study of the morphological characteristics of *Phoenix dactylifera* L. cultivars in Al-Madinah Al-Munawarah-Saudi Arabia. BMC Plant Biol. 2022;22(1):461.
- Cenci A. M Rouard 2017 Evolutionary analyses of GRAS transcription factors in Angiosperms. Front Plant Sci 8 273.
- Laroche J, Li P, Maggia L, Bousquet J. Molecular evolution of angiosperm mitochondrial introns and exons. Proc Natl Acad Sci USA. 1997;94(11):5722–7.
- Oghabian A, Greco D, Frilander MJ. IntERest: intron-exon retention estimator. BMC Bioinformatics. 2018;19(1):130.
- Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 2009, 37(Web Server issue):W202–208.
- Nystrom SL, McKay DJ. Memes: a motif analysis environment in R using tools from the MEME suite. PLoS Comput Biol. 2021;17(9):e1008991.
- Yang T, Liu R, Luo Y, Hu S, Wang D, Wang C, Pandey MK, Ge S, Xu Q, Li N, et al. Improved pea reference genome and pan-genome highlight genomic features and evolutionary characteristics. Nat Genet. 2022;54(10):1553–63.
- Mascagni F, Usai G, Cavallini A, Porceddu A. Structural characterization and duplication modes of pseudogenes in plants. Sci Rep. 2021;11(1):5292.
- Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R. Gene duplication as a major force in evolution. J Genet. 2013;92(1):155–61.



48. Ren Y, Ma R, Fan Y, Zhao B, Cheng P, Fan Y, Wang B. Genome-wide identification and expression analysis of the SPL transcription factor family and its response to abiotic stress in Quinoa (*Chenopodium quinoa*). *BMC Genomics*. 2022;23(1):773.
49. Dong H, Yan S, Jing Y, Yang R, Zhang Y, Zhou Y, Zhu Y, Sun J. MIR156-Targeted SPL9 is phosphorylated by SnRK2s and interacts with ABI5 to enhance ABA responses in *Arabidopsis*. *Front Plant Sci*. 2021;12:708573.
50. Bencivenga S, Simonini S, Benková E, Colombo L. The transcription factors BEL1 and SPL are required for cytokinin and Auxin Signaling during Ovule Development in *Arabidopsis*. *Plant Cell*. 2012;24(7):2886–97.
51. Qin M, Zhang Y, Yang Y, Miao C, Liu S. Seed-specific overexpression of SPL12 and IPA1 improves seed dormancy and grain size in Rice. *Front Plant Sci*. 2020;11:532771.
52. Santos CS, Carbas B, Castanho A, Vasconcelos MW, Vaz Patto MC, Domoney C, Brites C. Variation in pea (*Pisum sativum* L.) seed quality traits defined by Physicochemical Functional properties. *Foods* (Basel, Switzerland) 2019, 8(11).
53. Kwon S-J, Brown AF, Hu J, McGee R, Watt C, Kisha T, Timmerman-Vaughan G, Grusak M, McPhee KE, Coyne CJ. Genetic diversity, population structure and genome-wide marker-trait association analysis emphasizing seed nutrients of the USDA pea (*Pisum sativum* L.) core collection. *Genes Genomics*. 2012;34(3):305–20.
54. Kreplak J, Madoui MA, Cápál P, Novák P, Labadie K, Aubert G, Bayer PE, Gali KK, Syme RA, Main D, et al. A reference genome for pea provides insight into legume genome evolution. *Nat Genet*. 2019;51(9):1411–22.
55. Zhang X, Dou L, Pang C, Song M, Wei H, Fan S, Wang C, Yu S. Genomic organization, differential expression, and functional analysis of the SPL gene family in *Gossypium hirsutum*. *Mol Genet Genomics*. 2015;290(1):115–26.
56. Zhong H, Kong W, Gong Z, Fang X, Deng X, Liu C, Li Y. Evolutionary analyses reveal diverged patterns of SQUAMOSA promoter binding protein-like (SPL) Gene Family in *Oryza* Genus. *Front Plant Sci*. 2019;10:565.
57. Gandikota M, Birkenbihl RP, Höhmann S, Cardon GH, Saedler H, Huijser P. The miRNA156/157 recognition element in the 3'UTR of the *Arabidopsis* SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. *Plant Journal: Cell Mol Biology*. 2007;49(4):683–93.
58. Sánchez D, Ganfornina MD, Gutiérrez G, Marín A. Exon-intron structure and evolution of the lipocalin gene family. *Mol Biol Evol*. 2003;20(5):775–83.
59. Xu G, Guo C, Shan H, Kong H. Divergence of duplicate genes in exon–intron structure. 2012, 109(4):1187–92.
60. Sakharkar M, Long M, Tan TW, de Souza SJ. ExInt: an Exon/Intron database. *Nucleic Acids Res*. 2000;28(1):191–2.
61. Carvalho SG, Guerra-Sá R, de Merschmann C. The impact of sequence length and number of sequences on promoter prediction performance. *BMC Bioinformatics*. 2015;16(19):55.
62. Van Roey K, Gibson TJ, Davey NE. Motif switches: decision-making in cell regulation. *Curr Opin Struct Biol*. 2012;22(3):378–85.
63. Li J, Mahajan A, Tsai MD. Ankyrin repeat: a unique motif mediating protein-protein interactions. *Biochemistry*. 2006;45(51):15168–78.
64. Ciriello GD, Guerra CJBifg, proteomics: a review on models and algorithms for motif discovery in protein-protein interaction networks. 2008, 7 2:147–56.
65. Ren Y, Ma R, Xie M, Fan Y, Feng L, Chen L, Yang H, Wei X, Wang X, Liu K, et al. Genome-wide identification, phylogenetic and expression pattern analysis of HSF family genes in the Rye (*Secale cereale* L.). *BMC Plant Biol*. 2023;23(1):441.
66. Wang S, Wu K, Yuan Q, Liu X, Liu Z, Lin X, Zeng R, Zhu H, Dong G, Qian Q, et al. Control of grain size, shape and quality by OsSPL16 in rice. *Nat Genet*. 2012;44(8):950–4.
67. Si L, Chen J, Huang X, Gong H, Luo J, Hou Q, Zhou T, Lu T, Zhu J, Shang-guan Y, et al. OsSPL13 controls grain size in cultivated rice. *Nat Genet*. 2016;48(4):447–56.
68. Hu J, Huang L, Chen G, Liu H, Zhang Y, Zhang R, Zhang S, Liu J, Hu Q, Hu F, et al. The Elite alleles of OsSPL4 regulate grain size and increase grain yield in Rice. Volume 14. New York, NY: Rice; 2021. p. 90. 1.
69. Chao LM, Liu YQ, Chen DY, Xue XY, Mao YB, Chen XY. *Arabidopsis* transcription factors SPL1 and SPL12 Confer Plant Thermotolerance at Reproductive Stage. *Mol Plant*. 2017;10(5):735–48.
70. Zhao JL, Shi M, Yu J, Guo CK. SPL9 mediates freezing tolerance by directly regulating the expression of CBF2 in *Arabidopsis thaliana*. *BMC Plant Biol*. 2022;22(1):59.
71. Ma Y, Xue H, Zhang F, Jiang Q, Yang S, Yue PT, Wang F, Zhang YY, Li LG, He P, Zhang ZH. The miR156/SPL module regulates apple salt stress tolerance by activating MdWRKY100 expression. *Plant Biotechnol J*. 2021;19(2):311–23.
72. Li YX, Han SC, Sun XM, Ullah KN, Zhong Q, Zhang ZY, Zhang HL, Ming F, Li ZC, Li JJ. Variations in *OsSPL10* confer drought tolerance by directly regulating *OsNAC2* expression and ROS production in rice. *J Integr Plant Biol*. 2023;65(4):918–33.
73. Zhao Y, He JQ, Liu MM, Miao JN, Ma C, Feng YJ, Qian JJ, Li HH, Bi HH, Liu WX. The SPL transcription factor TaSPL6 negatively regulates drought stress response in wheat. *Plant Physiol Biochem*. 2024;206:108264.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.